

**What is claimed is:**

1. A method to create a fingerprint for media entities, comprising:  
 reading data indicative of a media entity desiring at least one fingerprint, said media entity data containing a sequence of random bits having a length N; and  
 processing said media entity data in accordance with at least one fingerprinting algorithm, said fingerprinting algorithm employing bit-to-bit comparisons and at least one approximation technique to process fingerprints;
2. The method as recited in claim 1, wherein said processing step further comprises the steps of:  
 calculating the average information density of said media entities;  
 determining the standard deviation of the calculated information of said media entities;  
 calculating the average critical band energy of the said media entities;  
 calculating the average standard deviation of the critical band energy of said media entities;  
 determining the play-time of said media entities; and  
 processing said information density, said standard deviation of said information density, said critical band energy, said standard deviation of said critical band, and said play time to produce a bit-sequence representative of said fingerprint.
3. The method as recited in claim 2, further comprising the step of comparing said bit sequence of said created fingerprint with said bit sequence of said data indicative of said media entities.
4. The method as recited in claim 3, wherein said comparing step contemplates the use of the Hamming distance between the fingerprint bit and the media entity bit to determine the probability that said fingerprint and said media entity bits differ by Hamming distance according to the relation,

$$P(M) = e^{-(M - N/2)^2 / 2\sigma^2} / \sigma\sqrt{2\pi},$$

wherein  $\sigma$  is the standard deviation of the distribution expressed as,

$$\sigma = \sqrt{N/2}.$$

5. The method as recited in claim 4, further comprising the step of calculating the probability that the Hamming distance between two sequences of random bits is less than a value  $M'$  according to the relation,

$$P(M < M') = \int_0^{M'-1} e^{-(x-N/2)^2/2\sigma^2} / \sigma \sqrt{2\pi} dx.$$

6. The method as recited in claim 2, wherein the average information density is taken to be the average entropy per processing frame of said media entities.

7. The method as recited in claim 6, wherein said average information density is determined by the relation,

$$S_{ave} = \frac{\sum_j S_j}{N}$$

wherein, N is the total number of processing frames.

8. The method as recited in claim 7, wherein  $S_j$  is determined by the relation,

$$S_j = -\sum_n b_n \log 2(b_n),$$

where  $b_n$  is the absolute value of the nth bin of the normalized real FFT of the processing frame.

9. The claim as recited in claim 8, where in the average standard deviation of the information density of said media entities is determined by the relation,

$$S_{std} = \frac{\sqrt{\sum_j (S_{ave} - S_j)^2}}{N}.$$

10. The method as recited in claim 2, wherein the average critical band energy is determined by the relation,

$$\vec{C}_{ave} = \frac{\sum_j \vec{C}_j}{N}$$

wherein,  $\vec{C}_j$  is a vector of values consisting of the critical band energy in each critical band and N is the total number of processing frames.

11. The method as recited in claim 2, wherein the average standard deviation of the critical band energy is determined by the relation,

$$C_{std} = \frac{\sqrt{\sum_j (C_{ave} - C_j)^2}}{N}$$

wherein, N is the total number of processing frames.

12. A computer readable medium bearing computer executable instructions for carrying out the method of claim 1.

13. A modulated data signal carrying computer executable instructions for carrying out the method of claim 1.

14. A computing device comprising means for carrying out each of the steps of the method of claim 1.

15. A system to create a fingerprint for media entities comprising:  
a sampling system;  
a processing system cooperating with said sampling system to generate said fingerprints, said processing system comprising one more means to calculate at least one of information density of said media entities, standard deviation of the information density of

said media entities, average critical band energy of said media entities, standard deviation of the critical band energy of said media entities, or the play-time of said media entities; and

a communications interface, said communications interface cooperating with said processing system to communicate created fingerprints to participating users.

5

16. The system as recited in claim 15, wherein said sampling system prepares at least one sampling portion of said media entities for communication to said processing system.

17. The system as recited in claim 16, wherein said processing system cooperates with said sampling system to process said sampling portion when generating said fingerprint.

10

18. The system as recited in claim 15, wherein said processing system comprises a computing environment capable of performing said calculations.

19. The system as recited in claim 18, wherein said computing environment comprises any of a stand-alone or networked computing environments.

20. The system as recited in claim 15, wherein said communications interface comprises any of a fixed-wire LAN, a wireless LAN, a fixed-wire WAN, a wireless WAN, a fixed-wire extranet, a wireless extranet, a fixed-wire intranet, a wireless intranet, peer-to-peer computer network, the wireless Internet, and the fixed-wire Internet.

21. The system as recited in claim 15, wherein said processing system is a component of a media content analysis and distribution system.

22. A method to identify media entities using fingerprints, comprising the steps of:  
calculating a fingerprint in accordance with the steps of claim 1 of said media entities;  
comparing said calculated fingerprint to already calculated fingerprints found in a cooperating fingerprint data store; and  
evaluating the results of the comparison.

30

09928001, 051001

23. The method as recited in claim 22, further comprising the step of communicating the results of said evaluation step to participating users, said participating users comprising any of: cooperating media entity processing systems, end-users, regulatory agencies.

24. A method to authenticate media entities to ensure compliance with copyright regulations by employing fingerprints, comprising the steps of:  
calculating a fingerprint in accordance the steps of claim 1 of said media entities;  
comparing said calculated fingerprint to fingerprints of authorized media entities stored in a cooperating data store; and  
evaluating the results of the comparison to return a response indicative whether authorization was granted.

25. The method as recited in claim 24, further comprising the step of denying distribution access to media entities that are determined to be unauthorized.

26. A method to calculate a fingerprint for media entities, comprising the steps of:  
reading a predefined amount of data from an input media entity data file  
corresponding to a specified position in said media entity data file;  
windowing said predefined amount of data into several sequential chunks;  
calculating the psycho-acoustic spectral coefficients;  
preserving the most energetic coefficients;  
calculating the inverse Discrete Fourier Transform (DFT) to generate an estimate of the salient coefficients;  
storing the results of the DFT for all said chunks into a matrix  $F$ , wherein said columns of said matrix  $F$  corresponds to a slice of time of said media entities and said rows of said matrix  $F$  correspond to a frequency band of the psycho-acoustic frequency scale;  
calculate the average of each row in said matrix  $F$ ;  
store results in vector  $\underline{E}$ ;  
calculate average of subset of elements in each row;  
store subset average in vector  $\underline{S}$ ;  
calculate vector  $\underline{D}$  such that  $\underline{D}$  is the difference between  $\underline{E}$  and  $\underline{S}$  and

quantizing each element in  $\underline{D}$  to a value of 1 if said each element value is greater than zero and to a value of 0 if said each element value is less than or equal to zero;

27. The method as recited in claim 26, further comprising the steps of:  
 5 assigning a fingerprint name to the calculated quantized vector  $\underline{D}$ ; and  
 storing said vector  $\underline{D}$  with said assigned fingerprint name in a cooperating fingerprint data store.

28. A method to identify media entities by employing media entity fingerprints,  
 10 comprising the steps of:  
 calculating the fingerprint for said media entities according to the steps of claim 26;  
 obtaining a sequence having length L of n random bits representing said calculated fingerprint;  
 obtaining a sequence having a length L of N random bits of said media entities desired  
 15 for identification;  
 comparing said n bits with said N bits; and  
 evaluating the results of said comparison.

29. The method as recited in claim 28, wherein said comparing step contemplates the use  
 20 of the Hamming distance between the fingerprint bit and the media entity bit to determine the probability that said fingerprint and said media entity bits differ by Hamming distance according to the relation,

$$P(M) = e^{-(M-N/2)^2/2\sigma^2} / \sigma\sqrt{2\pi},$$

wherein  $\sigma$  is the standard deviation of the distribution expressed as,

$$\sigma = \sqrt{N/2}.$$

30. The method as recited in claim 29, further comprising the step of calculating the probability that the Hamming distance between two sequences of random bits is less than a value M' according to the relation,

$$P(M < M') = \int_0^{M'-1} e^{-(x-N/2)^2/2\sigma^2} / \sigma\sqrt{2\pi} dx .$$

0923004-031001